Thuy Duong Nguyen

thuyduo3@ualberta.ca thuyduongnguyen thuyduongnguyen thuy-duong-nguyen

Personal Statement

My ultimate goal is to build AI agents that are supportive of and aligned with humans. To achieve this, I am interested in (i) understanding the reasoning and learning behaviors of contemporary Large Language Models (LLMs) and (ii) developing reliable, efficient agents that learn continually from evolving contexts and operate in interactive settings with humans.

Education

University of Alberta (UoA) Sep 2026 – Present
MSc in Computing Science (Thesis-based)

Hanoi University of Science and Technology Oct 2020 – Aug 2024
B.S. in Computer Science

- GPA: 3.73/4.00

Publications & Patents

[*] denotes equal contribution. Reverse-chronological order.

LongHarness Bench: Stress-testing Language Model Harnesses for Long-context Reasoning Preprint

Quang Hieu Pham*, ***Thuy Duong Nguyen***, Jocelyn Qiaochu Chen, Xi Ye

Retrieval Augmented Generation on Structured Data in Artificial Intelligence Models Filing US Patent

Dat Quoc Nguyen, Truong Dan Phan Dinh, ***Thuy Duong Nguyen***, Arvind Vardarajan Santhanam

ClozeMath: Improving Mathematical Reasoning in Language Models by Learning to Fill Equations ACL 2025 (Findings)

Quang Hieu Pham*, ***Thuy Duong Nguyen***, Tung Pham, Anh Tuan Luu, Dat Quoc Nguyen

Experience

AI Resident April 2025 – Aug 2026
Qualcomm AI Research

- Qualcomm acquired our VinAI's research team
- Focuses: Agentic AI, Large Language Models, Reasoning

AI Resident Aug 2024 – April 2025
VinAI Research

- Focuses: Large Language Models, Reasoning

Teaching Assistant Aug 2024 – Sep 2024
VinBigdata AI Engineer Training Program

- Courses: Machine Learning
- Prepare coding materials and participate in grading students' assignments

Viettel Digital Talent May 2024 – Aug 2024
Viettel AI & Data Services Center

- Focuses: Agentic AI, Retrieval-Augmented Generation

Projects

LongHarness Bench — University of Alberta

- Developed a 4-task benchmark of 200 problems testing diverse retrieval and long-context management strategies as well as their corresponding cost on state-of-the-art long-context agents.
- LongHarness is challenging. GPT-5.6-sol on the best harness tested only achieved $\approx 68\%$ accuracy.
- LongHarness reveals critical cost different between harnesses. Identical models can achieve similar performance across harnesses while incurring up to a $12.5\times$ difference in cost.
- No model-harness combination, across 25 tested combinations, consistently leads on every LongHarness's task.

E-commerce Agent — Qualcomm AI Research

- Refined a personal shopping database.
- Propose and implement a data curation and benchmark framework.
- Developed an FAQ module that achieved 95.5% accuracy on synthetic test set.
- Passed the Qualcomm global patent review, US patent application is currently filing.

Agent-UniRAG — Viettel AI & Data Services Center

- Designed an LLM agent framework for QA tasks.
- Distilled knowledge into small open-source LLMs, achieving performance parity with larger closed-source models across multiple RAG benchmarks.

Honors & Awards

University of Alberta Graduate Recruitment Scholarship	Fall 2026
Outstanding Resident Award in Applied Engineering	Aug 2026
Qualcomm Problem-Solvers Pack 2025 Award	Nov 2025
Top 0.02% in Vietnam's national university examination (Math-Physics-Chemistry)	Aug 2020

References

- **Asst. Prof. Jocelyn Chen** — University of Alberta
  MSc. supervisor
- **Asst. Prof. Xi Ye** — University of Alberta
  MSc. supervisor
- **Dr. Tung Pham** — Qualcomm AI Research
  Senior Researcher
- **Dr. Dat Quoc Nguyen** — Qualcomm AI Research
  Principal Researcher